Internet Threat Detection System Using Bayesian Estimation

Masaki Ishiguro[*1] Hironobu Suzuki[*2] Ichiro Murase[*1] Hiroyuki Ohno[*3]

Abstract. We present an Internet security threat detection system[*4] using Bayesian estimation method. This system analyzes security state of the Internet using Bayesian estimation with transition of frequencies of IP packet arrival events to some specified IP addresses such as port scanning, worm activities and so on. While the system calculates the frequency of access events in each time interval, Bayesian updating has been repeatedly applied to improve the confidence in degree of Internet critical states. When the system detects security threat(s) on the Internet, a security alert message is automatically sent to registered E-mail addresses, such as system administrators', and the system issues security alert details on our Web site[*5]. We also provide simple HTML and HDML for mobile phone browsers aka NTT DoCoMo's i-mode and KDDI's EZweb. Since the security state of the Internet changes dynamically, application of Bayesian estimation for threat detection is considered suitable because parameters of the model of Bayesian estimation are considered as dynamically changing quantities. This paper is focused on mechanism of detecting security threat using Bayesian estimation and our experimental evaluation.

1 Introduction

Recently, the number of large-scale security incidents has been increasing via network by computer worm (ex. MSblaster, Welchia) and malicious automatic scan/attack toolkit.

Time-lags between discovery of software vulnerabilities and distribution of this exploit codes is becoming shorter and shorter. Occasionally, 0-day attack is happened. This situation leads to serious damage for computer and network systems on the Internet.

Although many of intrusion detection systems (IDS) have been developed and operated, IDS only detects network attacks from evidence by signature and/or trace after intrusion into their network has occured. The large amount of IDS data and its statistical analysis are required if we want to know "what's going on the Internet". For example, Incidents.org provides InternetStormCenter (isc.incidents.org) and CERT/CC starts AirCERT project (www.cert.org). Still we need some way to find something from bunch of data.

Our goal is to provide "Open — 24 hours a day, 7 days a week" security threat watching service without human resources. So, we have developed a system that analyzes security state of the Internet using Bayesian estimation with transition of frequencies of IP packet arrival events to some specified IP addresses.

2 Internet Threat Detection Method

We describe an Internet threat detection method which detects critical state of the Internet caused by increase of wide area network attacks by analyzing time-series transition of frequencies of port scanning events to some specified IP addresses.

We define critical states of the Internet as "highly probable states where wide-area network attacks are becoming active in the Internet and may cause loss Tof srvices or damages to a network site". he Basic

[1]Information Technology Development Dept., Mitsubishi Research Institute, Inc

[2]Hironobu SUZUKI Office

[3]Emergency Communications Group, Communications Research Laboratory

[4]We call this "the Internet Weather Forecast System".

[5]`www.clscan.org`. This site is opened to the public.

idea of Bayesian estimation is to repeat updating a degree of confidence based on observations to gain improved probability distribution. Our method uses differences between frequency of port scans and their trend as observations.

Given a target time t (denoted as "estimation time" in Figure 1), we set some period of time (we call it "confidence update period").

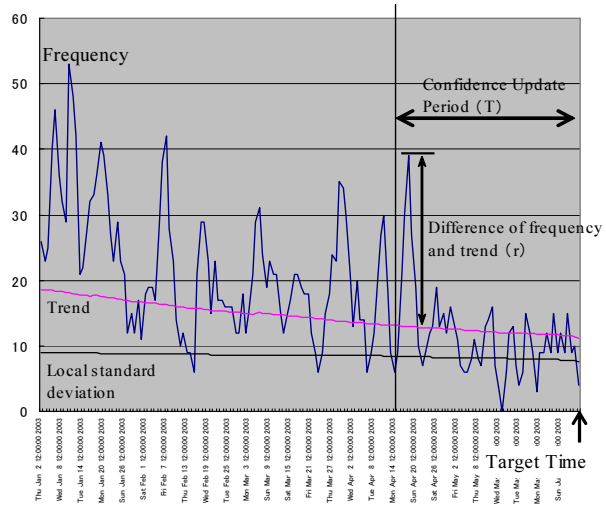

Figure 1: Time-series frequency of port scans and parameters of Bayesian estimation

We set up an initial prior probability distribution of confidence of critical states of the Internet. By observing frequency of port scans, moving average of frequency (we call it "frequency trend"), and standard deviation of difference between frequency and its trend at each time during confidence update period, we can update prior probability distribution to get posterior probability distribution using likelihood function defined in equation (1) and (2) based on updating equation (3).

$$P(r|s_0) = \frac{r}{k\sigma_r + r} \quad (1)$$

$$P(r|s_1) = \frac{k\sigma_r}{k\sigma_r + r} \quad (2)$$

$$P(s_i|r) = \frac{P(s_i)P(r|s_i)}{\sum_j P(s_j)P(r|s_j)} \quad (3)$$

where $s_i (i = 0, 1)$ represents states of the Internet as follows:

$$\begin{cases} s_0 & : \text{ critical state} \\ s_1 & : \text{ safe state} \end{cases} \quad (4)$$

r is observational value representing difference between frequency of port scans and their trend. $P(s_i|r)$ is a posterior probability that the Internet state is s_i after observing r. $P(s_i)$ is prior probability that the state is s_i before observing r. $P(r|s_j)$ is a likelihood function with which r is observed at state s_i. We defined likelihood functions (1) and (2) so that difference between frequency of port scans and their trend increases at critical state. In equations (1) and (2), σ_r is a local standard deviation of difference between frequency of port scans and their trend within the confidence updating period. k is a parameter defining sensitivity of updating speed (we call this "Update deviation factor"). Ranges of equations (1) and (2) are the real numbers between 0 to 1.

By repeating this update process at each time step within confidence updating period, we can get estimated confidence of critical status of the Internet.

3 An Overview of Internet Threat Detection System

An Internet threat detection system that we propose is to send alert messages to system administrators hopefully before their network sites have been attacked, by observing changes of port scanning to some specified IP addresses in the Internet and detecting threat caused by increasing activities of network attacks. Our system detects critical states of the Internet by analyzing time-series frequencies of port scans based on Bayesian estimation. Figure 2 is an overview of our Internet threat detection system and its environment.

Our system consists of a number of sensor systems, a log server, an analysis and detection system,

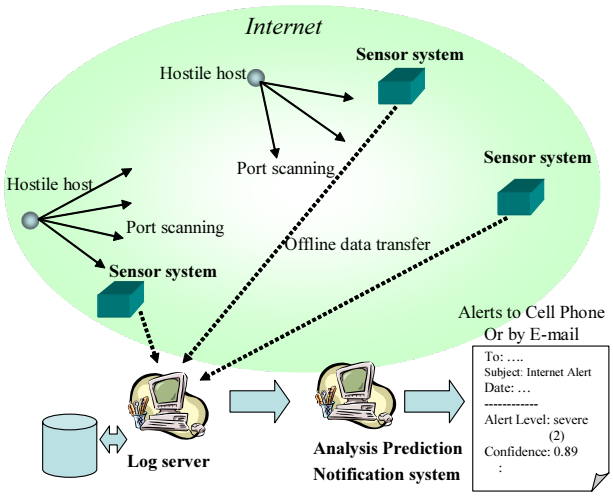

Figure 2: An overview of our Internet threat detection system

an notification system. The sensor systems record TCP/UDP access events making use of Linux syslog in clscan[4] compatible data format and sends them to the log server. The log server collects clscan data from several sensor systems and provides data to the analysis and detection system on demand. The analysis and detection system detects critical status of the Internet based on the method described in section 2 and request the notification system to send alerts. The notification system alerts system administrators about the critical status of the Internet. Currently, the notification system notifies critical status to system administrators by cell phone data communication system such as i-mode, ez web operated by cell phone companies in Japan.

4 Experimental Evaluation

We show experimental evaluation of our Internet threat detection system conducting by ROC (Receiver Operating Characteristics) analysis whose effectiveness is well-known in the field of signal detection research.

In ROC analysis, false-positive fraction which is a ratio that negative samples are mistakenly classified as positive samples, true-positive fraction which is a ratio that positive samples are correctly classified as positive are totally considered for performance evaluation. True-positive fraction represents sensitivity of detection. False-positive fraction corresponds to peculiarity of samples.

Our research aims at detecting critical states of the Internet where wide-area network attacks are increasing. It is difficult to evaluate performance of this type of problem, because true states of this kind of definition cannot be provided in advance in general. In this experiment, we specified manually the time during the sample period where the Internet is in critical state by observing time-series frequencies of port scans. We use this information as true state of the Internet and compares with the predication result obtained by using our system.

Data we used in this experiment is that in the period during January 1, 2003 to December 1, 2003. Figure 3 shows time-series transition of frequencies of port scans of some ports whose number of accesses are relatively large and on which JPCERT, an incident response organization, made critical alerts during this period.

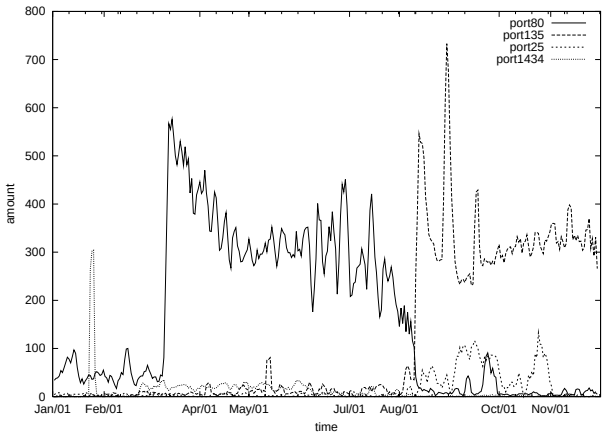

Figure 3: Time-series frequencies of port scans

Port 80 is used by http servers. An alert on Microsoft IIS 5.0 vulnerability of port 80 was made

on March 18, 2003 by JPCERT(JPCERT-AT-2003-0003). Port 135 is used by Windows RPC services. An alert on port scans to TCP port 135 which is well-known for W32/Blaster is made on August 15,2003 (JPCERT-AT-2003-0006). Port 25 is used by SMTP servers. An alert on sendmail vulnerability was made on March 3, 2003 (JPCERT-AT-2003-0004). Port 1434 is used by Microsoft SQL server 2000. An alert on port scans to UDP 1434 was made on January 27, 2003 (JPCERT-AT-2003-01-27).

Figure 4 is a time-series transition of estimation on degree of critical status of the Internet corresponding to Figure 3.

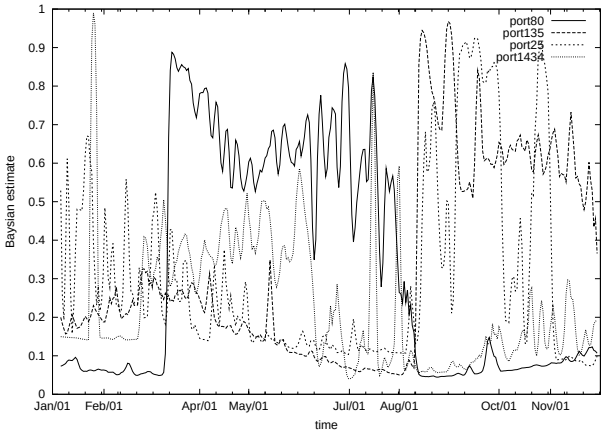

Figure 4: Time-series transition of Bayesian estimation on critical status

Among these ports we selected port 25 (SMTP) as a target of ROC analysis, because transition of frequency of port scans are relatively complex during this period and Warnings to this port were made by several incident response organizations such as CERT Advisory.

Figure 5 is a distribution of Bayesian estimation values at times in critical state (depicted as positive cases) and in safe state (depicted as negative cases).

We can see in Figure 5 that peak of distribution of positive cases is located at around 0.9 in horizontal axis and peak of distribution of negative cases is located at around 0.1, which we can confirm our

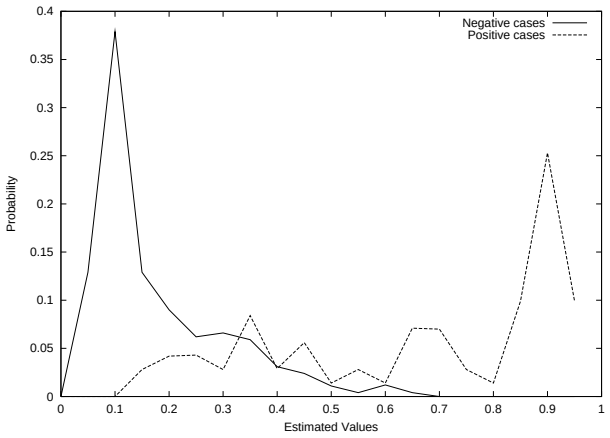

Figure 5: Distributions of Bayesian estimated values of critical and safe status

estimated values can be used as a good index for discrimination.

Figure 6 shows the ROC curve which can be obtained by plotting points composed of false-positive fraction and true-positive fraction. false-positive fraction can be calculated by summing up right-side area of positive cases from each estimated value in Figure 5 and true-positive fraction can be calculated by summing up negative cases.

Table tab:parameters shows experiments and their parameter values and performance value, that is a area size of ROC curve called Az values.

Table 1: Parameter values and Az Value for experiments

ID in Fig 6	Bayesian update deviation factor	Confidence update period	Trend period	Az
T301B5k0.5	0.5	5	301	0.95
T100B5k0.5	0.5	5	100	0.79
T100B3k0.5	0.5	3	100	0.80

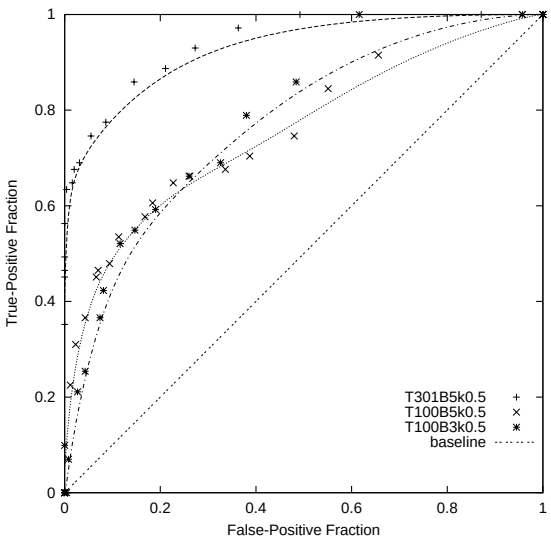

Figure 6: ROC curve of thread detection of port 25

ROC curves is located over the line $y = x$. The upper it is located, the better its performance is. In Figure6, we can conclude that the performance is better among the three experiment when trend period is 301 days, Bayesian update deviation factor is 0.5, and Confidence update period is 5 days.

5 Conclusions

We have developed a system which detects Internet threats caused by increasing activities of network attacks and makes alerts to system administrators hopefully before their network sites have been attacked. Internet threats are detected by observing time-series transition of frequencies of port scans to some specified IP addresses and applying Bayesian estimation. Degree of confidence in critical states on the Internet can be estimated by repeatedly updating the confidence based on the difference between frequency of port scans and its trend. Internet threats are automatically detected and alert messages are sent to system administrators.

By conducting ROC analysis, we have compared performance of detection and obtained a better parameter set of trend period, Bayesian update deviation factor, and Bayesian update period among three experiments.

References

[1] P. Porras and P. Neumann, "EMERALD: Event Monitoring Enabling Responses to Anomalous Live Disturbances",In Proceedings of the Nineteenth National Computer Security Conference, October 1997.

[2] Richard O. Duda et al.,Pattern Classification, John Wiley & Sons2001

[3] Masaki Ishiguro et al., "Discrimination Method of Malignant and Benign Tumors based on Machine Learning to Features of Liver CT images", Japanese Society of Medical Imaging Technology (JAMIT), Vol.19,No.1,pp.43-49.

[4] Hironobu Suzuki, clscan home page, http://www.pp.iij4u.or.jp/ h2np/h2np/lscan/

[5] Hitosi Mizutomo, "Development of DDoS Attack Predication System" IPSJ Computer Security Symposium 2003, pp.97-102.